600-76-0145

orig

# THE EFFECT OF CARRIER PRACTICES ON MEDICARE PREVAILING FEES: A SIMULATION

Jonah Otelsberg, Ph.D., Charlotte F. Muller, Ph.D., and Gertrude Oxenfeldt, M.A.*

Center for Social Research
City University of New York
Graduate Center
33 West 42 Street
New York, N.Y. 10036

Supported by DHEW/Health Care Financing Administration Contract #600-76-0145.

* Philip Drummond and David L. Miller provided Systems and Programming support.

American Public Health Association 107th Annual Meeting, 4 P.M., Wednesday, November 7, 1979, Statistics Section, Contributed Papers Session.

# Abstract

The payment basis for physicians' services under Medicare Part B is the reasonable charge concept, Medicare carriers have used different parameters and rules for deciding which claims to include in determining reasonable charges and how to group them by specialty and locality. Fees can be affected by these intercarrier differences. As part of a large-scale study of carrier practices, simulation methods were used to study effects of parameters defining extreme values, minimum numbers of services required to develop customary and prevailing charges, locality designation, specialty grouping, etc. The Wilcoxon Matched Pairs Signed Ranks Test was performed to test the significance of changes in practices on 50th and 75th percentile fees, the number of doctors included in computation of prevailings, and other measures. 1976 claims for 14 procedures processed by Group Health, Inc., carrier for Queens county in New York, were used in assessing the effects of parameter changes and the economic index. Nassau county claims processed by Blue Cross/Blue Shield of Greater New York were used to evaluate the effect of carrier programs. The parameters and programs used by GHI to derive reasonable charges served as a benchmark in the comparisons. The study showed that changed parameters altered the number of prevailings that could be formed and the number of doctors included. For various reasons, locality designation and other practices tested had little effect on prevailing fee levels but the economic index does affect fees.

Key Words: Medicare, Physician Fees, Wilcoxon Matched Pairs Signed Ranks Test

## Introduction

The payment basis for physicians' services under Medicare is the reasonable charge concept. Insurance carriers administering Part B of the Medicare program are directed by regulations and by the law itself to use this concept in setting fees for service. In pricing a given procedure, a carrier must derive for each physician a customary charge, defined as the median of submitted charges in a base year. Physicians are grouped by locality, and within locality, by specialty; for each such group an array of customaries, weighted by frequency of occurrences, is formed. The 75th percentile of the array is known as the unadjusted prevailing; it is adjusted for inflation control by application of an economic index. The payment basis for current Medicare claims is the lowest of: prevailing, customary, or submitted charge for a service.

Regulations leave considerable leeway to carriers in determining reasonable charges and the discretionary practices that have evolved can affect the cost of the program, access to services, and service quality. Carrier decisions include areas combined to form a Medicare locality, charges excluded from consideration in forming customaries and prevailings, specialty groups for which prevailings are determined separately, alternative pricing methods in absence of a customary or prevailing charge, and many others. Although use of reasonable charges is intended to make the Medicare market conform to the rest of the medical services market, i.e., to achieve "naturalistic" pricing of physician services to the elderly, actual outcomes of this process may vary among carrier territories.

Under a contract from the Health Care Financing Administration of the Department of Health, Education, and Welfare, the Center for Social Research at CUNY examined the effect of carrier practices in interpreting the non-mandated aspects of reasonable charge determination on Medicare fees. The study contained both national and local analyses of carrier discretionary practices and included relevant measures of administrative performance.

The present paper is drawn from the micro study. It analyzes the effects of alternative carrier practices in the reasonable charge process by comparing two carriers in contiguous counties in New York - Group Health, Incorporated in Queens and Blue Cross/Blue Shield of Greater New York in Nassau.

To assess the fee effect of carrier options in the reasonable charge determination process, a series of simulations was carried out based on claims for reimbursement for 14 commonly performed Medicare procedures. By processing the same claims set under identical conditions except for substitution of a single alternative assumption as to parameters or program rules, the effect of a discretionary practice is isolated from variations due to other causes. Alternative parameters for deriving customary and prevailing charges and various limits for defining extreme charges to be excluded from the process are tested. The effect on fees of different carrier approaches to locality formation, specialty grouping and merging of non-Medicare claims data from the carrier's private business and other sources are studied. In addition, the overall effect of a carrier's reasonable charge computer program is tested by putting claims submitted to one carrier through the program of the other carrier and comparing resulting fee levels to actual fees. Finally, as part of the simulation study, the effect of the Economic Index on prevailing fees is analyzed.

## Method

In studying the Medicare reasonable charge process in two neighboring counties, Queens (GHI) and Nassau (BC/BS-GNY), GHI reasonable charge programs were applied to claims for 14 selected procedures. The parameters used by

GHI served as a benchmark. The claims for which the prevailing charges were computed were selected from the list of 50 most frequently used procedures as reported in the Medical Directory of Prevailing Charges and are a representative sample of procedures for which GHI and Blue definitions were comparable. For each test of carrier practices, the number of possible comparisons was determined by available information for relevant procedure/specialty/type of service combinations. The type of service indicator is important when claims for the same procedure are submitted by a surgeon and an assistant surgeon both in the same specialty, and in other situations. The effect of the following carrier discretionary practices on reasonable charges was evaluated:

- changing the definition of extreme charges (outliers) excluded from computations
- changing the minimum number of charges needed to form a customary (the median of the distribution of submitted charges for a given physician for a given procedure within a calendar year.)
- changing the minimum number of customaries required to form a prevailing - the 75th percentile of the distribution of customaries, weighted by their frequencies, for a given procedure.
- merging Medicare claims with non-Medicare claims - claims from the carrier's own insurance business and any other claims, such as Title XIX (Medicaid) and Federal Employees Health Benefit Program (FEHBP) claims, which are included in the computation of prevailing charges.
- carrier computer programs
- locality designation - the division of the carrier's territory into geographic areas for which prevailing charges for Part B services will be derived.
- specialty grouping - practice fields used to group physician charges in computing prevailing charges. HCFA recognizes 44 areas of specialization in the Medicare Carriers Manual, but the carrier can use one, two, or more groups in the computation.
- Economic Index adjustment (The Economic Index is a price index by which prevailing charges used in determining Medicare reasonable charges for physicians' services may be increased above the level for FY 1973 only to the extent determined to be justified by the Secretary. The Economic Index figure is furnished by HCFA to all carriers each year.)

To test the significance of the effect of change in each practice, the Wilcoxon Matched Pairs Signed Ranks Test was used. This test was selected because it takes into account both the magnitude and the direction of the difference between two data sets. The measures on which the tests were performed were as follows:

- the 50th percentile of the weighted distribution of customaries
- the 75th percentile of the weighted distribution of customary charges (unadjusted prevailing)
- the number of doctors included in the computation of the prevailing charges
- the highest customary charge included in the computation of the prevailing charges
- the lowest customary charge included in the computation of the prevailing charges

Data bases for the tests varied somewhat according to the problem at hand. Tests of the effect of changing parameters used in computation of reasonable charges - extreme values, minimum needed to form a customary, and minimum needed to form a prevailing - were performed only on Medicare claims for CY (Calendar Year) 1976 supplied by GHI. The effect of inclusion of non-Medicare claims data in computation of reasonable charges was tested for GHI only using both Medicare and non-Medicare data for CY 1976 which GHI used to compute FSY (Fee Screen Year) 1978 prevailings. FSY is defined as the period July 1, year (M-1) to June 30, year (M).

To evaluate the overall effect of carrier computer programs, the BC/BS CY 1974 claims for Nassau were used. The FSY 1976 prevailings computed by the carrier were compared to those obtained by putting Blue CY 1974 claims through GHI programs. The effect of locality designation was tested by using claims provided by both carriers for CY 1976. Since BC/BS uses only Medicare claims to compute prevailings, only Medicare claims data of GHI were used in this comparison. For testing the effect of specialty grouping GHI Medicare claims data were used to develop prevailings for individual specialties, which were then compared to prevailings based on standard GHI specialty groupings. The influence of the economic index adjustment was tested by comparing FSY 1978 unadjusted and adjusted prevailings as provided by GHI.

## Effect of Parameter Changes on Prevailing Fees

In evaluating the effect of changes in parameters the prevailings based on GHI programs and parameters (as used by GHI to compute FSY 1978 prevailings) were considered the benchmark to which all the other computed prevailings were compared. Only the parameter being tested is changed in each test, while all the others remain as used by the carrier. The parameters used by GHI in computing prevailings were as follows:

- claims not within two standard deviations from the mean were excluded from the computation of each physician's customary ($S_2$)
- four charges were required to form a customary if all charges were different, and three if two were alike ($C_{4,3}$)
- five valid customaries were required to form a prevailing ($P_5$)

The following values of parameters were utilized in our tests:

Extreme values:

- values outside one standard deviation from the mean ($S_1$)
- values outside three standard deviations from the mean ($S_3$)

Number required to form a customary:

- one claim ($C_1$)
- five claims ($C_5$)
- ten claims ($C_{10}$)

Number required to form a prevailing:

- one customary ($P_1$)
- three customaries ($P_3$)
- nine customaries ($P_9$)

## Change in Definition of Extreme Values

### $S_1$ vs. $S_2$

Under the current parameters of two standard deviations from the mean used to exclude outlying charges, 80 combinations of procedure, specialty, and type of service were created in 1978 computations of prevailings by GHI.

If charges that lie outside one standard deviation from the mean are excluded, fewer prevailings can be created. Only 76 hypothetical prevailings could therefore be compared with the values that GHI actually developed (Table 1). Change to a one standard deviation parameter implies less toleration of extreme values in determining reasonable charges. There was no significant difference at either the 50th or the 75th percentile. However, statistically significant differences were found with regard to the number of doctors whose customamaries were used to compute the prevailing. Fewer doctors were used in 70% of the comparisons and in the remaining 30% there was no change. Another significant difference was in the highest charges included in the 76 computations; in more than half of the cases this charge was reduced. The result was similar for the lowest charges included. Outliers in both directions are affected by a tighter parameter but prevailings are not affected.

$S_3$ vs. $S_2$

Excluding charges that are further than three standard deviations from the mean does not change the number of prevailings that can be developed. This change implies more toleration of extreme values. No significant differences were seen at the 50th and 75th percentiles and the number of doctors whose fees were used in the prevailings was not affected by this change in parameter. Statistically significant differences were found in the comparison of the highest charges; in 23 of the 80 cases the charges were increased; in the remaining cases they were identical. The lowest charges included in the computation of prevailings were reduced in 13 cases by the change in parameters and identical in 67 cases. The looser parameter is more likely to bring in high values than low ones, which is to be expected in a positively skewed distribution.

### Change of Customary Parameters

$C_{01}$ vs. $C_{4,3}$

The customary charge is the median of submitted charges of a given physician for a procedure. Four observations were required to compute a valid customary (three, if two were equal). Reducing the required number of observations means that more customaries can be created and ultimately prevailings will be affected as well.

The number of prevailings computed when there are no restrictions on the number of claims needed to form a customary is 109 versus 80 when the benchmark parameters are used and some physicians do not have the minimum number of charges needed to form a customary (Table 2). No statistically significant differences were found with regard to the 50th and 75th percentile. There were more doctors whose customaries were included in the computation of the prevailing charge in 98% of the comparisons. Reducing the number of instances required would be expected

TABLE I

Change of Standard Deviation Parameters Compared to Benchmark a/

| | $S_1C_{4,3}P_5$ | $S_3C_{4,3}P_5$ |
|---|---|---|
| **50th Percentile** | | |
| Number of possible comparisons | 76 | 80 |
| Number of negative differences | 4 | 0 |
| Number of positive differences | 4 | 2 |
| Level of significance (Wilcoxon) | 0.3631 | 0.0899 |
| **75th Percentile** | | |
| Number of possible comparisons | 76 | 80 |
| Number of negative differences | 3 | 0 |
| Number of positive differences | 4 | 1 |
| Level of significance (Wilcoxon) | 0.3999 | 0.1587 |
| **Number of Doctors** | | |
| Number of possible comparisons | 76 | 80 |
| Number of negative differences | 53 | 0 |
| Number of positive differences | 0 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | N.A. |
| **High Customary Charge** | | |
| Number of possible comparisons | 76 | 80 |
| Number of negative differences | 39 | 0 |
| Number of positive differences | 0 | 23 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0000* |
| **Low Customary Charge** | | |
| Number of possible comparisons | 76 | 80 |
| Number of negative differences | 0 | 13 |
| Number of positive differences | 42 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0007 |
| Number of cases for Benchmark $S_2C_{4,3}P_5$ | 80 | 80 |
| Number of cases for test | 76 | 80 |

a/ $S_2$ - standard deviations

$C_{4,3}$ - Minimum of 4 charges or 3 charges (if two are alike) to form a customar

$P_5$ - Minimum of 5 valid customaries to form a prevailing.

* This figure is significant at a level beyond 4 decimal places.

N.A.: Not applicable as there are no differences to compare.

to bring in more doctors with extreme customary charges. In fact, in almost half of the cases, the highest charge was raised by this change in parameter, but the lowest charges were reduced in only 1/5 of the 80 comparisons. The implication is that the looser parameter shifts the distribution of customaries towards higher values more often than towards lower values.

### $C_{05}$ vs. $C_{4,3}$

When five claims per doctor are necessary to form a customary, fewer prevailings are computed - 64 versus 80 for the benchmark. Significant differences are found on all measures but the 75th percentile. The 50th percentiles are lower in 4 cases and the same in all the rest. There were fewer doctors whose customaries were used to produce the prevailing in 61 out of the 64 possible comparisons. The highest charges are reduced in 7 out of 64 comparisons, while the lowest charges are raised in the same proportion of cases, indicating a tighter distribution of customaries.

### $C_{10}$ vs. $C_{4,3}$

When the minimum number of charges required for computing a customary is increased, fewer customaries are produced and there is a possibility that fewer prevailings can be created. In this simulation only 54 hypothetical prevailings ultimately were created that could be compared with prevailings developed using existing parameters.

For the 50th percentile there is a statistically significant difference, as 9 of the 54 comparisons had lower prevailings. Fewer differences occurred at the 75th percentile comparison. In almost all comparisons the number of doctors whose customaries were used in computing the prevailing was significantly less. The highest customary in the distribution was reduced in almost 1/2 of the cases, but the lowest charge was raised in about 1/4 of the cases and not

TABLE II

Change of Customary Parameters Compared to Benchmark a/

| | $S_2C_{01}P_5$ | $S_2C_{05}P_5$ | $S_2C_{10}P_5$ |
|---|---|---|---|
| **50th Percentile** | | | |
| Number of possible comparisons | 80 | 64 | 54 |
| Number of negative differences | 3 | 4 | 9 |
| Number of positive differences | 2 | 0 | 1 |
| Level of significance (Wilcoxon) | 0.1726 | 0.0339 | 0.0264 |
| **75th Percentile** | | | |
| Number of possible comparisons | 80 | 64 | 54 |
| Number of negative differences | 1 | 0 | 4 |
| Number of positive differences | 3 | 0 | 0 |
| Level of significance (Wilcoxon) | 0.1367 | N.A. | 0.0339 |
| **Number of Doctors** | | | |
| Number of possible comparisons | 80 | 64 | 54 |
| Number of negative differences | 0 | 61 | 52 |
| Number of positive differences | 78 | 0 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0000* | 0.0000* |
| **High Customary Charge** | | | |
| Number of possible comparisons | 80 | 64 | 54 |
| Number of negative differences | 0 | 7 | 24 |
| Number of positive differences | 37 | 0 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0090 | 0.0000* |
| **Low Customary Charge** | | | |
| Number of possible comparisons | 80 | 64 | 54 |
| Number of negative differences | 16 | 0 | 0 |
| Number of positive differences | 0 | 7 | 13 |
| Level of significance (Wilcoxon) | 0.0002 | 0.0090 | 0.0007 |
| Number of cases for Benchmark $S_2\ C_{4,3}P_5$ | 80 | 80 | 80 |
| Number of cases for test | 109 | 64 | 54 |

a/

$S_2$ - 2 standard deviations

$C_3$ - Minimum of 4 charges or 3 charges if 2 are alike to form a customary

$P_5$ - Minimum of 5 valid customaries to form a prevailing.

* This figure is significant at a level beyond 4 decimal places.

N.A. - Not applicable as there are no differences to compare.

affected in the rest. The implication is that doctors infrequently performing a Medicare procedure are more likely to charge "above market" than below.

Change of Prevailing Parameters

$P_1$ vs. $P_5$, $P_3$ vs. $P_5$, $P_9$ vs. $P_5$

Changing the number of valid customaries used to compute the prevailing charges had no effect on the 50th and 75th percentile fees (Table 3). Of course, the number of doctors represented changes, going down as more customaries are required. The current standard calls for five customaries. The most important effect of changing the number of customaries required is on the number of prevailings computed. It changes from 126 when only one customary is required to compute a prevailing to 92 when three are required, 80 when five, and 59 when nine customaries are required. This means that as more customaries are required to form a prevailing, fewer prevailings will be computed by the reasonable charge process and supplementary methods have to be developed to complete the price structure. This involves higher administrative costs.

Discussion of Effect of Parameter Changes

The change in parameter defining extreme values does not have a significant effect on either the 50th percentile or the prevailing, but the number of doctors is significantly affected.

When the minimum number for customaries is reduced to one (no minimum) or increased to five, the value of prevailings computed is not greatly affected; only when the minimum is increased to 10 is there a significant effect on the value of prevailings. The 50th percentile is much more sensitive to the change. The number of prevailings that can be computed is greatly affected by the parameter; as the minimum increases, i.e., as more claims are required

TABLE III

## Change of Prevailing Parameters Compared to Benchmark [a/]

| | $S_2C_{4,3}P_1$ | $S_2C_{4,3}P_3$ | $S_2C_{4,3}P_9$ |
|---|---|---|---|
| **50th Percentile** | | | |
| Number of possible comparisons | 80 | 80 | 59 |
| Number of negative differences | 0 | 0 | 0 |
| Number of positive differences | 0 | 0 | 0 |
| Level of significance (Wilcoxon) | N.A. | N.A. | N.A. |
| **75th Percentile** | | | |
| Number of possible comparisons | 80 | 80 | 59 |
| Number of negative differences | 0 | 0 | 0 |
| Number of positive differences | 0 | 0 | 0 |
| Level of significance (Wilcoxon) | N.A. | N.A. | N.A. |
| **Number of Doctors** | | | |
| Number of possible comparisons | 80 | 80 | 59 |
| Number of negative differences | 0 | 0 | 0 |
| Number of positive differences | 0 | 0 | 0 |
| Level of significance (Wilcoxon) | N.A. | N.A. | N.A. |
| **High Customary Charge** | | | |
| Number of possible comparisons | 80 | 80 | 59 |
| Number of negative differences | 0 | 0 | 0 |
| Number of positive differences | 0 | 0 | 0 |
| Level of significance (Wilcoxon) | N.A. | N.A. | N.A. |
| **Low Customary Charge** | | | |
| Number of possible comparisons | 80 | 80 | 59 |
| Number of negative differences | 0 | 0 | 0 |
| Number of positive differences | 0 | 0 | 0 |
| Level of significance (Wilcoxon) | N.A. | N.A. | N.A. |
| Number of cases for Benchmark $S_2C_{4,3}P_5$ | 80 | 80 | 80 |
| Number of cases for test | 126 | 92 | 59 |

[a/] $S_2$ - 2 standard deviations

$C_{4,3}$ - Minimum of 4 charges, or 3 charges if 2 are alike, to form a customary.

$P_5$ - Minimum of 5 valid customaries to form a prevailing.

N.A. - Not applicable as there are no differences to compare.

to form a customary, the number of prevailings goes down from 109 for no minimum, to 80 for 4, 64 for 5, and 54 for 10 claims. As fewer procedures will have a prevailing computed by the reasonable charge program, supplementary means of obtaining prevailings are needed.

The number of doctors with valid customaries is significantly modified by the change in parameter, as is to be expected, and goes down as the required minimum increases. This implies that as the minimum increases, more doctors are outside the standard reasonable charge determination process. The extent to which this would occur would vary among procedures.

As more claims are required to form a customary, the distribution of customaries is tighter, i.e., the extreme values are closer to the measures of central tendency.

Changing the number of customaries required to form a prevailing affects only how many prevailings can be computed. As the minimum number of customaries goes up the number of prevailings goes down, with the result that an auxiliary method is needed for pricing Medicare claims.

## Effect of Merging Non-Medicare Charges

Carriers are directed to merge or compare Medicare claims with claims from their non-Medicare business. Not all carriers are in a position to do this effectively because of the nature of their private business. GHI-New York is one of the carriers that merges its private insurance claims with Medicare claims to compute prevailings, but the number of non-Medicare claims included is about a tenth of the number of Medicare claims. Not all of GHI's non-Medicare charges are included by GHI in its computation because of differences in coding and processing of the claims. In our simulation, prevailings computed for Medicare claims and non-Medicare claims were compared to each other and to the prevailings as computed by GHI using both sets of claims. The same computer programs and parameters were used throughout the computations.

Medicare vs. Non-Medicare

The more claims there are in a file, the larger the probable number of "cases" (prevailing charges that met the criterion for the minimum number of valid customaries) in the file. There were 190 prevailing charges computed for Medicare data and 72 for non-Medicare data. The number of possible comparisons for all measures of interest (cases where valid prevailings were derived in both files) was 68. For the 50th percentile of the weighted customaries, the two files are not significantly different, but for the 75th percentile they are. The non-Medicare claims data produced higher prevailing charges in 26 out of 68 comparisons whereas only 6 non-Medicare prevailings were lower than Medicare.

The most significant difference between the files is in the number of doctors whose customaries were used to compute the prevailings. In 58 of the 68 comparisons, there were fewer doctors for non-Medicare than for Medicare prevailings. In 8 cases there were more non-Medicare doctors. This again results from file size, since there are fewer doctors in the non-Medicare file.

When the highest charges for the two files are compared, the non-Medicare value is higher in 42 out of the 68 possible cases and lower in 21 cases. When the magnitude of the differences is taken into consideration, such a result could be expected by chance only 19 out of 10,000 times - i.e., it is highly significant. The "low charges" in the non-Medicare file were higher in 28 out of the 68 comparisons and lower in 20. When the size of the differences is taken into account the results are significant at the 1% level. Such a result

## TABLE IV

### Comparison of Non-Medicare to Medicare

### 50th Percentile

| | |
|---|---|
| Number of possible comparisons | 68 |
| Number of negative differences** | 13 |
| Number of positive differences | 20 |
| Level of significance (Wilcoxon) | 0.2777 |

### 75th Percentile

| | |
|---|---|
| Number of possible comparisons | 68 |
| Number of negative differences | 6 |
| Number of positive differences | 26 |
| Level of significance (Wilcoxon) | 0.0001* |

### Number of Doctors

| | |
|---|---|
| Number of possible comparisons | 68 |
| Number of negative differences | 58 |
| Number of positive differences | 8 |
| Level of significance (Wilcoxon) | 0.0000* |

### High Customary Charge

| | |
|---|---|
| Number of possible comparisons | 68 |
| Number of negative differences | 21 |
| Number of positive differences | 42 |
| Level of significance (Wilcoxon) | 0.0019* |

### Low Customary Charge

| | |
|---|---|
| Number of possible comparisons | 68 |
| Number of negative differences | 20 |
| Number of positive differences | 28 |
| Level of significance (Wilcoxon) | 0.0114* |

| | |
|---|---|
| Number of cases for Benchmark (Medicare) | 190 |
| Number of cases for test (Non-Medicare) | 72 |

---

* Significant at $P \leq .01$

** Non-Medicare less than Medicare

could be expected by chance 114 out of 10,000 times.

Medicare vs. Total

The number of cases is 198 for the total file (including Medicare and non-Medicare records) and 190 for the Medicare file; the number of possible comparisons for all measures of interest (valid prevailings in both files) was 190 (Table 5).

Although there are more negative differences (Medicare smaller than total) than positive differences, when the amount of the difference is taken into account the two files do not differ significantly with regard to the 50th percentile of the distribution of the weighted customaries, and there is practically no difference between the two files with regard to the 75th percentile.

The two files differ significantly as to the number of doctors considered in the comparison of the prevailings, the Medicare file including fewer doctors than the total file in about half of the cases.

When the highest charges are compared, Medicare charges are lower in 53 cases, higher in 2 cases and identical in 135 cases. The differences are statistically significant. An opposite picture emerges when the lowest charges are compared. Medicare charges are higher in 30 cases and the remaining 160 cases have identical charges.

Non-Medicare vs. Total

There are 72 cases in the non-Medicare file compared to 198 in the total file (Table 5). The number of possible comparisons is limited by the number of cases in the smaller file and, therefore, is equal to 72.

Although the differences between the files as to the 50th percentile of the weighted customaries are not statistically significant, when the

TABLE V

Comparison of Medicare and Non-Medicare to Total

| | 50th Percentile | Medicare | Non-Medicare |
|---|---|---|---|
| Number of possible comparisons | | 190 | 72 |
| Number of negative differences | | 10 | 14 |
| Number of positive differences | | 4 | 18 |
| Level of significance (Wilcoxon) | | 0.1813 | 0.3003 |
| | 75th Percentile | | |
| Number of possible comparisons | | 190 | 72 |
| Number of negative differences | | 5 | 6 |
| Number of positive differences | | 2 | 22 |
| Level of significance (Wilcoxon) | | 0.2495 | 0.0006 |
| | Number of Doctors | | |
| Number of possible comparisons | | 190 | 72 |
| Number of negative differences | | 97 | 70 |
| Number of positive differences | | 0 | 0 |
| Level of significance (Wilcoxon) | | 0.0000 | 0.000 |
| | High Customary Charge | | |
| Number of possible comparisons | | 190 | 72 |
| Number of negative differences | | 53 | 26 |
| Number of positive differences | | 2 | 27 |
| Level of significance (Wilcoxon) | | 0.0000 | 0.4912 |
| | Low Customary Charge | | |
| Number of possible comparisons | | 190 | 72 |
| Number of negative differences | | 0 | 6 |
| Number of positive differences | | 30 | 33 |
| Level of significance (Wilcoxon) | | 0.000 | 0.0 |
| Number of cases for Benchmark (Total) | | 198 | 198 |
| Number of cases for test (Medicare, Non-Medicare) | | 190 | 72 |

75th percentiles are compared the non-Medicare is higher in 22 cases, lower in 6 cases and equal to total in 44 cases. These differences are significant. The most striking difference between the two files is in the number of doctors. In 70 out of 72 cases, the non-Medicare file had fewer doctors whose customaries were used to compute the prevailings. This, like the number of cases, results from file size. The two files do not differ significantly as to the high charges but the low charges in the non-Medicare file are significantly higher than those in the total file.

Discussion of Effect of Merging Non-Medicare Charges

While there are significant differences between the prevailings computed from each of the files separately, they do not affect the prevailings used by GHI because the non-Medicare file is relatively small. The central tendencies of the two files are similar. The 75th percentiles differ because there are outliers in the non-Medicare file. However, if more non-Medicare claims were included, they might significantly affect the reasonable charge.

## Effect of Carrier Computer Programs

The effect of carrier computer programs, including the parameters used, on prevailings was evaluated by comparing prevailings for FSY 1976 computed by BC/BS for Nassau to those computed from the same claims using GHI programs. (BC/BS CY 1974 claims were converted into GHI format prior to the computation). This comparison is limited to prevailing charges since no other information was available in the BC/BS record.

The two systems differ in definition of extreme values and specialty groupings. BC/BS excludes values lower than 1/3 or higher than three times the current charge, but GHI uses two standard deviations to define the limits.

BC/BS computes prevailings for general practitioners and specialists - two categories, whereas GHI uses more specialty groupings. Other differences undoubtedly exist, but without access to BC/BS programs these differences could not be studied in detail. The comparison of the prevailings, it is believed, provides an overall comparison of the computer systems used in reasonable charge determination.

Our results (Table 6) show no significant differences between the prevailings obtained from the two systems. Of the 38 possible comparisons, 30 showed a difference; in 13 cases the BC/BS prevailing was lower, in 17 it was higher.

### Effect of Locality Designation

In order to test the effect of locality designation on prevailings the CY 1976 Medicare claims of both carriers were combined and FSY 1978 prevailings were computed using GHI programs and parameters. These prevailings were compared to those computed using GHI programs on Medicare claims for Queens and Nassau separately. For the 75th and the 50th percentiles, no significant difference is found between the combined file and each of its components (Table 7). The number of doctors whose customaries were used in computation of prevailings is significantly higher in the combined file than in the components, as is to be expected. The conclusion, then, is that combining these two counties into one locality would not significantly change the prevailing charges in these counties.

### Effect of Specialty Groupings

To test the effect of specialty groupings, the prevailing fees computed for Medicare claims using the standard GHI groupings were compared to prevailings computed for each specialty individually. The groupings used by GHI

TABLE VI

## Comparison of BC/BS Computed* Prevailing Charges to Those Provided by BC/BS

| | |
|---|---|
| Number of possible comparisons | 38 |
| Number of negative differences | 13 |
| Number of positive differences | 17 |
| Level of significance (Wilcoxon) | 0.2964 |
| | |
| Number of cases for computed BC/BS prevailings | 80 |
| Number of cases for provided BC/BS prevailings | 69 |

* BC/BS prevailings were computed using GHI programs with the parameters $S_2C_3P_3$ which are as close as we could get to BC/BS parameters.

$S_2$ - charges outside two standard deviations from the mean were excluded from computation of customaries.

$C_3$ - three charges were needed to form a valid customary.

$P_3$ - three customaries were needed to form a valid prevailing.

were as follows: Family Practice (08), comprising specialties in Family Practice (08) and Pediatrics (37); Internal Medicine (11), comprising specialties in Internal Medicine, (11), Allergy (03), Cardiovascular Diseases (06), Gastroenterology (10), and Pulmonary Diseases (29); and the third grouping, Surgical Specialties (02), which includes Surgery (02), Neurological Surgery (14), Plastic Surgery (24), Proctology (28), and Thoracic Surgery (33).

For purposes of comparison, each specialty included in the group was assigned the group prevailing (benchmark), which was compared to the prevailing computed for that specialty (test). The number of cases for the benchmark was 156 and the number of possible comparisons was 89 (Table 8). The 75th percentile was not significantly affected by specialty designation.

### Effect of Economic Index Adjustment

The effect of the Economic Index adjustment on the 50th and 75th percentiles was evaluated by comparing GHI adjusted to unadjusted charges, FSY 1978. There were 67 possible comparisons for both measures and results showed significant differences in that the adjusted charges were lower at the 50th percentile in more than 1/4 of the cases, and at the 75th percentile in more than half the cases (Table 9). Since the adjusted 75th percentile represents the administrative upper bound for reasonable charges (subject to the "no roll-back" provision), it seems that the Economic Index restrains program outlay.

### Summary

When parameters used in the computation of prevailings are changed, the two items that respond most often are the number of doctors on whose customaries the prevailings were based and the number of prevailings computed. The actual

TABLE VII

[illegible] Medicare ([illegible]) and BC/BS ([illegible]) [illegible] Prevailing

| | CHI | BC/BS |
|---|---|---|
| | **50th Percentile** | |
| Number of possible comparisons | 80 | 79 |
| Number of negative differences | 16 | 12 |
| Number of positive differences | 16 | 10 |
| Level of significance (Wilcoxon) | 0.5128 | 0.2697 |
| | **75th Percentile** | |
| Number of possible comparisons | 80 | 79 |
| Number of negative differences | 11 | 16 |
| Number of positive differences | 8 | 12 |
| Level of significance (Wilcoxon) | 0.9039 | 0.2845 |
| | **Number of Doctors** | |
| Number of possible comparisons | 80 | 79 |
| Number of negative differences | 61 | 61 |
| Number of positive differences | 0 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0000* |
| | **High Customary Charge** | |
| Number of possible comparisons | 80 | 79 |
| Number of negative differences | 35 | 20 |
| Number of positive differences | 0 | 0 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0001 |
| | **Low Customary Charge** | |
| Number of possible comparisons | 80 | 79 |
| Number of negative differences | 0 | 0 |
| Number of positive differences | 38 | 14 |
| Level of significance (Wilcoxon) | 0.0000* | 0.0010 |
| Number of cases for Benchmark 1/ | 106 | 106 |
| Number of cases for test | 80 | 79 |

1/ Combined CHI and BC/BS.

* This figure is significant at a level beyond four decimal places.

TABLE VIII

## Comparison of Prevailing Charges - Grouped vs. Individual Specialties

| | |
|---|---|
| Number of possible comparisons | 89 |
| Number of negative differences | 4 |
| Number of positive differences | 4 |
| Level of significance (Wilcoxon) | 0.8886 |

| | |
|---|---|
| Number of cases for Benchmark | 156 |
| Number of cases for test | 89 |

TABLE IX

## Comparison of GHI Adjusted Prevailing Charges to Unadjusted Prevailings, FSY 1978

| | 50th Percentile |
|---|---|
| Number of possible comparisons | 67 |
| Number of negative differences | 19 |
| Number of positive differences | -- |
| Level of significance (Wilcoxon) | 0.0001 |

| | 75th Percentile |
|---|---|
| Number of possible comparisons | 67 |
| Number of negative differences | 38 |
| Number of positive differences | -- |
| Level of significance (Wilcoxon) | 0.0000* |

| | |
|---|---|
| Number of cases for Benchmark | 67 |
| Number of cases for test | 82 |

* This figure is significant at a level beyond four decimal places.

values of the 50th and 75th percentiles are fairly resistant to the changes examined. When fewer doctors have customaries and fewer procedures have prevailings, the cost of administering the program increases since alternative methods of determining reasonable charges have to be employed. In addition, these methods are less clearly prescribed by HCFA than the reasonable charge process itself and thus the span of carrier discretion (leading to differences among carriers) is increased. Paradoxically, then, mandating higher minima for customary and prevailing fees may actually increase carrier discretion and variation in fees and other measures.

The effect of non-Medicare charges on prevailings for GHI is negligible mainly because of the small number of non-Medicare claims involved. Using one carrier computer program rather than another, in the case of GHI and BC/BS, does not have a significant effect on the level of the prevailings even though these carriers differ in the parameters used to determine extreme values and in specialty grouping. Combining Nassau and Queens into one locality did not affect the size of the prevailings, nor did a change in specialty groups. However, adjustment by the Economic Index does lower the charges significantly.

Simulation trials involving actual carrier data and using the Wilcoxon Matched Pairs Signed Ranks Test are helpful in showing how Medicare administrative rules and parameters interact with real market conditions (i.e., how many physicians actually perform given procedures, and how often) to affect reasonable charge levels and the number of physicians covered by the principal pricing method. Policy-makers can take the demonstrated effects into account in modifying the Medicare method of pricing physicians' services and in using Medicare as a model for national health care.

CMS LIBRARY
3 8095 00005775 8